Genetic Variation and Molecular Darwinism

Werner Arber
University of Basel, Department of Microbiology, Biozentrum, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

Meyers: Encyclopedia of Molecular Cell Biology and Molecular Medicine: Systems Biology,
Second Edition. Edited by Robert A. Meyers.

Keywords

Biological evolution
A nondirected, dynamic process of diversification resulting from the steady interplay between spontaneous mutagenesis and natural selection.

DNA rearrangement
Results from mostly enzyme-mediated recombination processes, which can be either intramolecular or intermolecular.

Evolution gene
The protein product of this gene acts as a generator of genetic variations and/or as a modulator of the frequency of genetic variations.

Gene acquisition
A result of the horizontal transfer of genetic information from a donor cell to a receptor cell. With bacteria this can occur in transformation, conjugation or phage-mediated transduction.

Natural selection
A result of the capacity of living organisms to cope with the encountered physico-chemical and biological environments. Largely depending on its genetic set-up and physiological phenotype, each organism may have either a selective advantage or a selective disadvantage as compared to the other organisms present in the same ecosystem.

Spontaneous mutation
Defined as any alteration of nucleotide sequences occurring to DNA, without the intended intervention of an investigator. The term *mutation* is used here as a synonym of genetic variation.

Transposition
DNA rearrangement mediated by a mobile genetic element such as a bacterial insertion sequence (IS) element or a transposon.

Variation generator
An enzyme or enzyme system, the mutagenic activity in the generation of genetic variation has been documented.

The comparison of DNA sequences of genes and entire genomes offers interesting insights into the possible evolutionary relatedness of the genetic information of living organisms. Together with a relatively rich database from experimental microbial genetics, conclusions can be drawn on the molecular mechanisms by which genetic variations are spontaneously generated. A number of different specific mechanisms contribute to the overall mutagenesis. Here, these mechanisms are grouped into three natural strategies of the spontaneous generation of genetic variations: local changes of DNA sequences; the intragenomic rearrangement of DNA segments; and the acquisition of foreign DNA by horizontal gene transfer. These three strategies have different qualities with regards to their contributions to the evolutionary process. As a general rule, none of the known mechanisms producing genetic variants is clearly directed. Rather, the resulting alterations in the inherited genomes are more random. In addition, usually only a minority of resulting variants provide a selective advantage. Interestingly, in most of the molecular mechanisms involved, the products of so-called ''evolution genes'' are involved as generators of genetic variation and/or as modulators of the frequencies of genetic variation. The products of evolution genes operate in tight collaboration with nongenetic factors, such as the structural flexibilities and chemical instabilities of molecules, chemical and physical mutagens, and random encounters. All of these aspects, by contributing to the spontaneous generation of genetic variations, together form the core of the theory of molecular evolution, which brings the Neo-Darwinism to molecular Darwinism. The philosophical and practical inferences of this knowledge will be discussed.

1 Introduction

Traditionally, evolutionary biology has devoted its major attention to the comparison of phenotypical traits of higher organisms, both of those actually living and of those that have become extinct (e.g., paleontological fossils). The resulting theory of descent is, together with other criteria, at the basis of the systematic classification of living organisms. Darwin's theory of natural selection brought a new element into the understanding of the long-term development of forms of life. Natural selection is the result of the coping of the organisms with the encountered living conditions which are dependent on both the environmental physico-chemical conditions and the activities of all living forms in a particular ecological niche. The Darwinian theory of evolution also postulated that intrinsic properties of life are not entirely stable and principally identical for all organisms of a given species. In the so-called modern synthesis, in which evolutionary biology and genetics became integrated, transmissible

phenotypic variations representing the substrate of natural selection were explained as being due to genetic variations (or mutations). Shortly thereafter, deoxyribonucleic acid (DNA) was identified as the carrier of genetic information, and DNA is therefore also the target for mutagenesis. Within the past few decades, a rapid development of molecular genetics with novel research strategies leading to genomics, the sequencing of entire genomes and functional studies of genes and their products, has paved the way to hitherto inaccessible knowledge on the basis of life and its multiple manifestations. This also relates to the process of molecular evolution. A synoptical insight into the various molecular mechanisms contributing to the generation of genetic variations represents a molecular synthesis between the Neo-Darwinian theory and molecular genetics. This synthesis toward molecular Darwinism can confirm the Darwinian evolution at the molecular level.

2
Principles of Molecular Evolution

The principles of molecular evolution to be outlined here are founded on:

1. The Neo-Darwinian theory of evolution, with its three pillars of genetic variation, natural selection, and isolation.
2. The solidly established microbial genetics database.
3. DNA sequence comparisons with bioinformatic tools.
4. Physico-chemical knowledge of the reactivity, the conformational flexibility, and the chemical stability of biologically active molecules.

2.1
Evolutionary Roles of Genetic Variation, Natural Selection, and Isolation

The long-term maintenance of any form of life requires a relatively high stability of its genetic information. However, rare, occasional genetic variations occur in all organisms, and these give rise to mixed populations of organisms with the parental genome and organisms with one or more alterations in their genome. Such populations are steadily submitted to natural selection, yet experience shows that – in general – favorable genetic variations are considerably less frequent than unfavorable variations, which provide a selective disadvantage. Indeed, genetic variants with unfavorable genetic alterations will, sooner or later, be eliminated from propagating populations, which will become enriched with organisms carrying favorable genetic variations. It should be noted that, by far, not all alterations in the nucleotide sequences of a genome will lead to a change in the phenotype of the organism. However, such silent and neutral mutations may later become physiologically relevant in conjunction with still other, upcoming DNA sequence changes.

Natural selection is by no means a constant element; rather, it varies both in time and in space due to variations not only in the physico-chemical environmental conditions but also in the life activities of the many different organisms that are present in a particular ecological niche and which form an ecosystem. Since a genetic variation may of course also affect the influence that the organism exerts on the other organisms present in the same ecosystem (consider weeds and pathogenicity effects, as well as beneficial, synergistic effects), any novel mutation may influence not only the life of the

concerned organism itself but also the lives of other cohabitants of the same ecological niche.

The third pillar – besides genetic variation and natural selection – of biological evolution is isolation, which evolutionary biologists define as two different types. The first type is geographic isolation, which may seriously reduce the number of potential habitats for an organism, while the second type is termed *reproductive isolation*. For example, two distantly related diploid organisms may not be fertile in sexual reproduction. However, reproductive isolation can also be seen in a wider definition to seriously limit the possibility of the horizontal transfer of segments of genetic information between two different types of organism.

As summarized in Fig. 1, genetic variation drives biological evolution; indeed, a complete genetic stability would render any evolutionary process impossible. A very high frequency of genetic variation would rapidly lead to the extinction of the concerned organisms because of the stated prevalence of unfavorable mutations in the spontaneous generation of genetic variations. It is natural selection, together with the available sets of genetic variants, which determines the direction of biological evolution or, in other words, the directions in which the branches of the ''tree of evolution'' grow. Finally, the geographic and reproductive isolations modulate the evolutionary process.

2.2 Molecular Mechanisms of the Generation of Genetic Variation

The concept to be presented here requires the reader to question some long-established textbook knowledge, such as the claim that spontaneous mutations would largely result from errors, accidents, and illegitimate processes. We defend here the alternative view that living nature actively cares for biological evolution:

1. by making use of intrinsic properties of matter such as a certain degree of chemical instability and of structural flexibility of molecules; and
2. by having developed genetically encoded systems, the products of which are involved in the generation of genetic variations and in modulating the frequencies of genetic variation.

It might be relevant to mention here that the term *mutation* is defined differently in classical genetics and in molecular genetics. In classical genetics, a mutant is any variant of a parental form showing in its phenotypic properties some alteration that becomes transmitted to the progeny. In contrast, it has become a habit in molecular genetics to call any alteration in the parental nucleotide sequence of the genome a mutation, whether it has phenotypic consequences or not. There is good reason to assume that, in most spontaneously occurring mutagenesis events, the specific mechanism involved will not pay attention to whether the sequence alteration at the involved target site will cause a phenotypic alteration or not. For studies on mechanisms and on the statistics of their occurrence, it is thus indicated to follow the molecular genetics definition of the term mutation, which we use here as a synonym of genetic variation. The term ''*spontaneous mutation*'' will be used to label any type of DNA sequence alteration unintended by the investigator. Yet, this definition says nothing about whether a mutation relates to a phenotypic change.

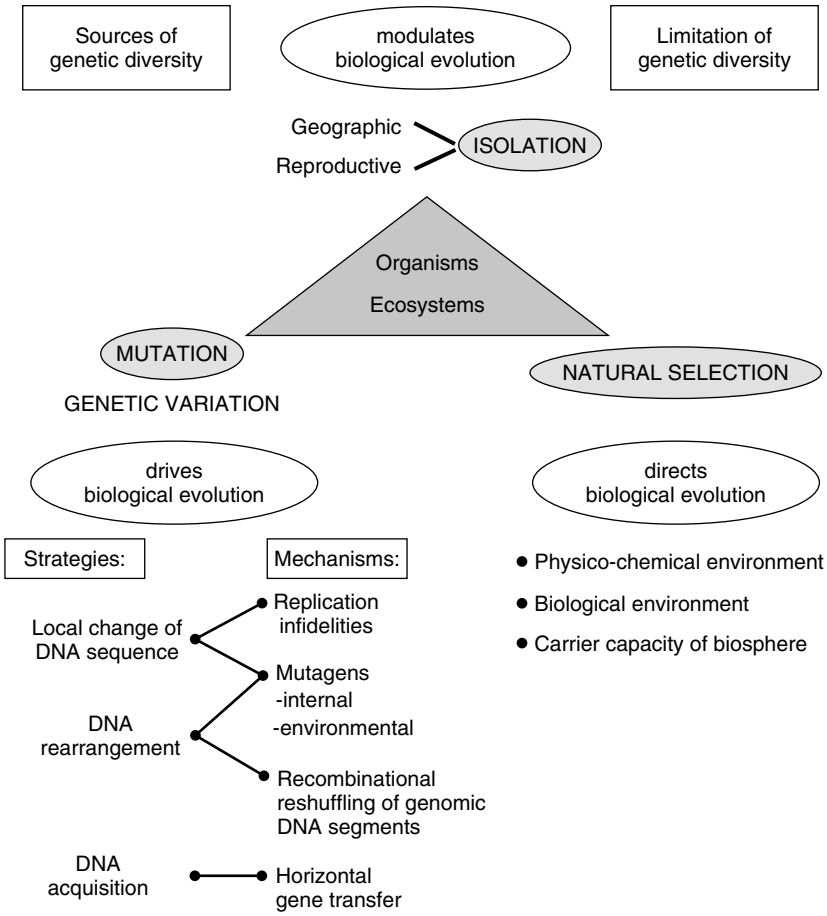

Fig. 1 Synoptical presentation of major elements of the theory of molecular evolution. A number of specific mechanisms contribute, each with its own characteristics to the four groups of mechanisms of genetic variation listed. Each of the specific mechanisms follows one (and sometimes more than one) of the three principal, qualitatively different strategies of genetic variation.

At present, research investigations into genetic variation follows two main strategies:

1. The increasing availability of entire and partial genome sequences offers excellent means to compare regulatory sequence elements, specific domains of genes, entire genes, groups of genes, and entire genomes with regard to DNA sequence homologies and genome organization. Within a given species, this can reveal a genetic polymorphism, whilst between more-or-less related species it can provide a reliable measure of the evolutionary relatedness. For example, the molecular clock – an indicator of evolutionary relatedness – is based on single nucleotide alterations. Sequence comparisons can often suggest how sequence alterations might have

occurred during the course of past evolution.
2. A more reliable insight into the generation of genetic variations can be gained by the observation of individual processes of mutagenesis. In view of the large size of genomes, and of the rare and random occurrence of spontaneous DNA sequence alterations, this approach is relatively difficult. However, a rich database is already available from microbial genetics, particularly from bacteria and their viruses and plasmids. Their relatively small genomes are haploid, so that phenotypic effects caused by genetic variation normally become rapidly manifest. With appropriate selection and screening techniques this can allow the identification of occasional, functionally relevant mutations in populations. On the other hand, investigations into structural alterations in the genomes of individual bacterial colonies – for example, by the study of restriction fragment length polymorphism (RFLP) – can reveal when and where on the genome a DNA rearrangement must have occurred.

A quite solid, general result of this type of experimental investigation reveals that, in the spontaneous generation of genetic variations, it is not only a single mechanism that is at work. Rather, a number of mechanistically different processes contribute to the overall mutagenesis, and selected examples will be discussed below. Interestingly, a critical evaluation of the situation shows that the specific mechanisms of mutagenesis often depend both on nongenetic elements and on the activities of specific enzymes – the products of so-called ''evolution genes.'' The multitude of the thus-identified, distinct mechanisms which contribute to the formation of genetic variations can be grouped into three qualitatively different natural strategies (Fig. 1):

1. Local changes in the DNA sequences.
2. The rearrangement of DNA segments within the genome.
3. The acquisition by horizontal transfer of a DNA segment originating in another type of organism.

Selected examples for each of these strategies will be discussed in Sections 4–6.

3 Genetic Variation in Bacteria

Several seminal discoveries, based largely on investigations carried out with microorganisms between 1943 and 1953, were essential for the subsequent development of molecular genetics:

1. It was realized that bacteria and bacteriophages have genes that can mutate and recombine, and that spontaneous mutations in microorganisms normally arise independently of the presence of selective agents [1, 2]. It was also learned that the genetic information of bacteria, and of some bacteriophages, is carried in DNA molecules rather than in other biological macromolecules, such as proteins [3].
2. The newly discovered phenomena of DNA transformation, bacterial conjugation and bacteriophage-mediated transduction revealed natural means of horizontal gene transfer between different bacterial cells [4].
3. It was seen that horizontal gene transfer has natural limits, including systems of host-controlled modification,

which today are known as *DNA restriction-modification systems* [5].

4. Mobile genetic elements were identified as sources of genetic instability, and were seen to represent mediators of genetic rearrangements [6]. While such rearrangements are often caused by transposition, alterations in the genome can also result from the integration of a bacteriophage genome into the genome of its bacterial host strain, which is thereby rendered lysogenic.

It was at the end of this fruitful period of discovery, in 1953, that the structural analysis of DNA molecules led to the double-helix model [7]. The filamentous structure of DNA molecules made it clear how genetic information could be contained in the linear sequences of nucleotides. The double helical nature of the model also offered an understanding of semiconservative DNA replication at the molecular level, and thus of information transfer into progeny.

Many classical microbial genetic investigations were carried out with *Escherichia coli* K-12, the genome of which is a single circular DNA molecule (chromosome) of about 4.6×10^6 base pairs (bp). In periods of growth, the rate of spontaneous mutagenesis is about 10^{-9} per bp and per generation, which represents one new mutation in every few hundred cells in each generation. Today, *E. coli* has several well-studied bacteriophages and plasmids, and these materials continue to facilitate many investigations into life processes in these bacteria.

Under good growth conditions, the generation time of *E. coli*, measured between one cell division and the next, is very short, on the order of 30 min. Upon exponential growth, this leads to a multiplication factor of 1000 every 5 h; thus, a population of 10^9 cells, representing 30 generations, is reached from an inoculum of a single cell in only 15 h. This rapid growth rate greatly facilitates population genetic studies and, in turn, investigations into the evolutionary process.

On the filiform DNA molecules of *E. coli* and its bacteriophages and plasmids, the genetic information is relatively densely stored as linear sequences of nucleotides or base pairs. Genes depend on the presence of continuous sequences of base pairs (reading frames) that encode specific gene products, usually proteins, and of expression control signals that ensure the occurrence of gene expression at the relevant time with the required efficiency. Mutations can affect reading frames as well as control signals, both of which represent specific DNA sequences. In addition, some specific DNA sequences relate to the control of the metabolism of the DNA molecules and, in particular, of their replication.

In the following sections, selected examples will be described of mechanisms that contribute, each in its specific manner, to the formation of genetic variants. The examples will be grouped into the three above-mentioned major natural strategies (Fig. 1) although, as will be seen later, some of the specific mechanisms involve more than one of these strategies.

4 Local Changes in the DNA Sequences

The process of DNA replication is one of the important sources of genetic variation, and may depend both on the structural features of the substrate DNA and on the functional characteristics of the replication fork [8, 9].

Some of the ''infidelities'' of DNA replication are likely to depend on tautomeric forms of nucleotides – that is, a structural flexibility which is inherent to these organic compounds [10]. Base-pairing depends on specific structural forms, and conformational changes of nucleotides can result in a mispairing if short-lived, unstable tautomeric forms are ''correctly'' used in the synthesis of the new complementary strand upon DNA replication. For this reason, mutations that result from this process are considered not as errors but rather as ''infidelities.'' DNA replication is, indeed, one of the sources of nucleotide substitution, and this plays an important role in the evolutionary development of biological functions.

Another source of mutagenesis is represented by an inherently low degree of chemical instability of the nucleotides. For example, cytosine can undergo oxidative deamination to become uracil; upon replication, this gives rise to an altered base pairing and results also in nucleotide substitution.

Other replication infidelities that may relate to slippage in the replication fork can result in either the deletion or the insertion of one or a few nucleotides in the newly synthesized DNA strand. If such mutations occur within reading frames for protein synthesis, the phenotypic effect may be drastic. Indeed this is the case when, in the protein synthesized from a gene affected by a frameshift mutation, the amino acid sequence downstream of the site of mutation differs greatly from that of the nonmutated product. In addition, the size of such a mutated protein is usually altered, depending on the chance occurrence of an appropriate stop codon in the new reading frame.

Proofreading devices and other enzymatic repair systems prevent replication infidelities from producing mutations at intolerably high rates. Generally, these act rapidly after replication by screening for imperfect base-pairing in the double helix. Successful repair thereby requires a specific means to distinguish the newly replicated DNA strand from its template – the complementary parental strand. Because of these correction activities, many primary mispairings are removed before they have the opportunity to become fixed as mutations. As a consequence, DNA repair systems are able to modulate the frequencies of mutagenesis.

The genetic information of some viruses – and also, on occasion, segments of genetic information of chromosomal origin – may pass through RNA molecules, which may later become retrotranscribed into DNA. To date, no efficient repair systems are known to act at the level of RNA; indeed, RNA replication shows a higher degree of infidelity than DNA replication. In consequence, genetic information that becomes replicated as RNA molecules generally suffers increased mutation rates.

A relatively large number of internal and environmental chemicals exert mutagenic effects by means of molecular mechanisms that, in many cases, are well understood and often cause local DNA sequence changes. Some intermediate products of the normal metabolic activities of a cell may be mutagenic, and may thus contribute to spontaneous local mutagenesis. The mutagens of the environment include a multitude of chemical compounds, as well as ultraviolet radiation and some physico-chemical constraints such as elevated temperature, by influencing the chemical stability of the nucleotides. Each of these mutagens and mutagenic conditions contributes in a specific way to the generation of genetic variations.

Again, many of the potential sequence alterations brought about by internal and environmental mutagens are efficiently repaired by enzymatic systems. Since, however, the efficiency of such repair is rarely 100%, evolutionarily relevant mutations persist.

5
Intragenomic DNA Rearrangements

Various recombination processes are well known to mediate DNA rearrangements, which often result in new nucleotide sequences.

Whilst in haploid organisms, general recombination is not essential for propagation, it does influence genetic stability at the population level in various ways as a generator of new sequence varieties. For example, it can bring about sequence duplications and deletions by acting at segments of homology that are carried at different locations in a genome. General recombination is also known to act in the reparation of damage caused to DNA by ionizing radiation. In this case, an intact genome can be assembled from the undamaged segments of sister copies of the chromosome by homologous recombination. The best known contribution of general recombination to genetic diversity is meiotic recombination, bringing about random recombinants between the paternal and the maternal chromosomes in diploid organisms [11].

Two other widely spread types of recombination system are described in greater detail in Sections 5.1 and 5.2, namely site-specific recombination and transposition, both of which are known to contribute to genetic variation. Still other recombination processes, such as that mediated by DNA gyrase, can – at least for the time being – be best grouped as ''illegitimate recombinations.'' This group may contain several different molecular mechanisms that act with very low efficiency and have remained, at least in part, unexplained.

5.1
Site-Specific DNA Inversion at Secondary Crossover Sites

Genetic fusions represent the results of joining together segments of two genes (gene fusions) or of two operons (operon fusions) that are not normally together. An operon is a set of often functionally related genes that are copied into messenger RNA (i.e., transcribed) as a single unit. As a result of this organization, these genes are coordinately regulated – that is, they are turned on or off at the same time. Therefore, in an operon fusion one or more genes are put under a different transcription control, although the genes *per se* remain unchanged. In contrast, gene fusion results in a hybrid gene composed of sequence motifs and often functional domains originating in different genes.

In site-specific DNA inversion, a DNA segment bordered by specific DNA sequences acting as sites of crossing over becomes periodically inverted by the action of the enzyme DNA invertase [12]. Depending on the location of the crossover sites, DNA inversion can give rise to gene fusion or to operon fusion. The underlying flip-flop system can result in microbial populations composed of organisms with different phenotypic appearances: if, for example, the DNA inversion affects the specificity of phage tail fibers, as is the case with phages P1 and μ of *E. coli*, phage populations with two or even more different host ranges will result [13, 14].

Occasionally, a DNA sequence that deviates considerably from the efficiently used

crossover site – a so-called ''secondary crossover site'' – can serve in DNA inversion, which thus involves a normal crossover site and a secondary crossover site [15, 16]. This process results in novel DNA arrangements, many of which may not be maintained because of lethal consequences or reduced fitness; however, if a few new sequences are beneficial for the life of the organism, these may be selectively favored. This DNA rearrangement activity can thus be regarded as evolutionarily important. As many different DNA sequences can serve in this process as secondary crossover sites (though at quite low frequencies), site-specific DNA inversion systems may act as variation generators in large populations of microorganisms. Thus, it has been postulated that this evolutionary role of DNA inversion systems may be more important than their much more efficient flip-flop mechanism, which can at most help a microbial population to more readily adapt to two different, frequently encountered environmental conditions. In fact, other strategies could also be used for the latter purpose.

Computer-aided comparisons of DNA sequences quite often reveal that independent genes may consist of a particular domain with high homology and of other DNA sequences which show no significant signs of relatedness. DNA inversion using secondary sites of crossing over is a potential source of such mosaic genes. DNA inversion can span over relatively large distances in DNA molecules [17], and has the advantage of not losing any DNA sequences located between the two sites of crossing over. Although deletion formation represents another source for gene fusion, it has the disadvantage that the DNA sequences between the sites of crossing over are usually eliminated.

5.2 Transposition of Mobile Genetic Elements

Nine different mobile genetic elements have been found to reside, often in several copies, in the chromosome of *E. coli* K-12 derivatives. This adds up to an occupation of about 1% of the chromosomal length by such insertion sequences (ISs; also termed *IS elements*). At rates on the order of 10^{-6} per individual IS element and per cell generation, these mobile genetic elements undergo transpositional DNA rearrangements. These include the simple transposition of an element and more complex DNA rearrangements such as DNA inversion, deletion formation, and the cointegration of two DNA molecules [6, 18]. Because of different degrees of specificity in the target selection upon transposition, the IS-mediated DNA rearrangements are neither strictly reproducible nor fully random. Transposition activities thus also act as variation generators. In addition to DNA rearrangements mediated by the enzyme transposase, which is usually encoded by the mobile DNA element itself, other DNA rearrangements simply take advantage of extended segments of DNA homologies at the sites of residence of identical IS elements, at which general recombination can act. Altogether, IS elements represent a major source of genetic plasticity for microorganisms.

Transposition occurs not only in growing populations of bacteria, but also during prolonged phases of rest. This effect is readily seen with bacterial cultures stored at room temperature in stabs (small vials containing a small volume of growth medium in agar). The stabs are inoculated with a drop of a bacterial culture taken up with a platinum loop, which is inserted (''stabbed'') from the top to the bottom of the agar. After an overnight incubation,

the stab is tightly sealed and stored at room temperature. Most strains of *E. coli* are viable in stabs during several decades of storage. That IS elements exert transpositional activities under these storage conditions is easily seen, as follows.

A stab can be opened at any time, a small portion of the bacterial culture removed, and the bacteria well suspended in liquid medium. After appropriate dilution, the bacteria are spread on a solid medium and grown overnight. Individual colonies are then isolated. The DNA from such subclones is extracted and fragmented with a restriction enzyme; the DNA fragments are then separated using gel electrophoresis. Subsequently, Southern hybridization with appropriate probes can be used to determine whether different subclones reveal restriction fragment length polymorphisms (RFLP), which are indicative of the occurrence of mutations during storage.

If this method is applied to subclones isolated from old stab cultures, and if DNA sequences from residential IS elements serve as hybridization probes, an extensive polymorphism is revealed. However, no or only minimal polymorphism occurs with hybridization probes from unique chromosomal genes. Good evidence is available that transposition represents a major source of this genomic plasticity observed in stabs, which at most allow for a very residual growth at the expense of dead cells. Thus, it can be concluded that the enzymes promoting transposition are steadily present in the stored stabs. Indeed, the IS-related polymorphism increases linearly with time of storage for periods of up to 30 years. In fact, in a culture of *E. coli* strain W3110 analyzed after 30 years of storage, each surviving subclone had suffered on average about a dozen RFLP changes, as identified with hybridization probes from eight different residential IS elements, of which IS5 was the most active. Lethal mutations could, of course, not be identified in this study [19–21].

Lethal mutations that affect essential genes for bacteriophage reproduction can be accumulated in the prophage state of the phage genome in its lysogenic host. Such mutants can be screened for their inability to produce infective phage particles upon induction of phage reproduction. Experiments were carried out with *E. coli* lysogenic for a phage P1 derivative grown in batch cultures at 30 °C for about 100 generations, allowing for alternative periods of growth and rest. Most of the independent lethal mutants could thereby be identified as being caused by the transposition of an IS element from the host chromosome into the P1 prophage that is maintained in its host as a plasmid. In these experiments, IS2 was the most active element, and was mainly inserted into a few preferred regions of the P1 genome, but each time at another site [22]. The insertion targets used did not show any detectable homology or similarity with each other. This is another good example of an enzymatically mediated variation generator, since the experiment as such identifies IS transposition as a major source of lethal mutagenesis [16].

Currently, there is no evidence available that bacterial mobile genetic elements would play an essential role in the bacterial life span extending from one cell division to the next. However, these elements are major players in the evolution of bacterial populations; as seen here, they clearly contribute to intragenomic DNA rearrangements. Depending on the target sequences involved, the resulting mutations may often be lethal by interrupting essential reading frames or expression control regions. Although favorable mutations may be relatively rare, they can

contribute to evolutionarily advantageous developments of the genome. That mobile genetic elements also play important roles in the natural strategy of DNA acquisition will be detailed in Section 6.

6 DNA Acquisition

While the mutagenesis mechanisms belonging to the strategies of local changes in the DNA sequences (Section 4) and of intragenomic DNA rearrangements (Section 5) are exerted within the genome and can affect any part of the genome, an additional strategy of spontaneous sequence alterations depends on an external source of genetic information. In DNA or gene acquisition, genetic information indeed originates from an organism other than that undergoing mutagenesis [23]. In bacteria, DNA acquisition can occur by means of transformation, conjugation, or virus-mediated transduction. In the latter two strategies of horizontal gene transfer, a plasmid or a viral genome, respectively, acts as natural gene vector.

The association and dissociation of chromosomal genes with natural gene vectors often arises from transpositional activities, and from general recombination acting at IS elements that are at different chromosomal locations. These mechanisms have been well studied, with conjugative plasmids and bacteriophage genomes serving in specialized transduction. For example, composite transposons – which are defined as two identical IS elements flanking a segment of genomic DNA (often with more than one gene unrelated to the transposition process) – are known occasionally to transpose into a natural gene vector and, after their transfer into a receptor cell, to transpose again into the receptor chromosome. Hence, together with other mechanisms – such as site-specific and general recombination – transposition represents an important promoter of horizontal gene transfer.

Several natural factors seriously limit gene acquisition. Transformation, conjugation, and transduction depend on the surface compatibilities of the bacteria involved. Furthermore, when the donor DNA penetrates into receptor cells it is very often confronted with restriction endonucleases; these enzymes cause fragmentation of the invading foreign DNA, which is subsequently completely degraded [5]. Before fragments become degraded, however, they are recombinogenic and may find the chance to incorporate all or part of their genetic information into the host genome. Therefore, the role of restriction systems can be interpreted as follows: they keep the rate of DNA acquisition low, whilst at the same time stimulating the fixation of relatively small segments of acquired DNA to the receptor genome. This strategy of acquisition in small steps can best offer microbial populations the chance to occasionally extend their biological capacities, without any extensive risk of disturbing the functional harmony of the receptor cell by acquiring too many different functions at once. These considerations have their relevance at the level of selection exerted on the hybrids resulting from horizontal DNA transfer. In fact, this selection represents one of the last steps in the acquisition process.

DNA acquisition by horizontal gene transfer is a particularly interesting source of new genetic information for the receptor bacterium, because the chance that the acquired DNA will exert useful biological functions is quite high. Most likely, it has already assumed the same functions in the donor bacterium.

An interesting hypothesis links the universality of the genetic code with the important role played by horizontal gene transfer in the evolutionary development of the living world [24]. According to this view, those organisms which use the most common genetic language would in the long term be able to profit best from the increasing worldwide pool of genetic functions, under pressure to adapt to changes in their living conditions.

7 The Three Natural Strategies Generating Genetic Variations Contribute Differently to the Evolutionary Process

Biological evolution is a systemic process [25] and, as outlined above, many different and specific mechanisms contribute to generate genetic diversity that represents, at any time, the substrate for natural selection. The building-up of functional complexity is a stepwise process, in which many random attempts at genetic alteration become rapidly rejected, while relatively few novel sequences are approved as favorable by natural selection and are maintained and amplified. The genome can thus be seen as a ''cabinet'' in which key informations from favorable historical developments are stored. Stepwise, additional favorable information is added. In the context of changing selective conditions, stored information having lost its functionally beneficial relevance may be deleted or favorably altered. As has been seen, a multitude of different mechanisms are behind this dynamic process, and for a better understanding of events the identified mechanisms have been grouped into three major natural strategies of genetic variation – local changes in the DNA sequences, intragenomic DNA rearrangements, and DNA acquisition – all of which have different qualities with regard to their contributions to biological evolution.

The local DNA sequence change is probably the most frequently involved strategy of genetic variation. Indeed, its frequency – which depends primarily on intrinsic properties of matter, chemical instability, and conformational flexibility – would be intolerably high if it could not be modulated by the efficient enzymatic systems of DNA repair. Local sequence changes bring about nucleotide substitution, the deletion, and insertion of one or a few base pairs, or a local scrambling of a few base pairs. These sequence changes can contribute to a stepwise improvement of a biological function. It must be borne in mind that the functional test for such improvement is carried out by natural selection. Alternatively, a long series of stepwise local sequence changes could be expected to bring about a novel biological function. However, this type of long-term process can achieve high efficiency only when natural selection begins to be exerted on such an upcoming function.

In contrast, the reshuffling of DNA segments within the genome can be considered as ''tinkering'' with the existing elements, whereby favorable gene fusions and operon fusions may occasionally result. DNA rearrangement may also be the source of gene duplication and higher amplification, both of which are widely recognized contributions to the evolutionary progress.

The evolutionarily high effectiveness of horizontal gene transfer has already been highlighted in Section 6. As a matter of fact, DNA acquisition allows the recipient organism to share in the success of evolutionary developments made by others.

In drawing the evolutionary tree of bacteria, DNA acquisition should be accounted for by more or less randomly adding temporal horizontal shunts between individual branches [16]. It must be borne in mind that, usually, only small DNA segments flow through such shunts in horizontal gene transfer.

Strictly speaking, several of the specific mechanisms of genetic variation employ more than one of the three natural strategies described above. In the transposition of IS elements, for example, a chromosomal DNA segment consisting of the mobile genetic element can undergo a translocation and thereby become inserted at a new target site. As a rule, the target sequence will as a consequence be duplicated, and this usually involves a few nucleotides. Hence, this transposition event will consist of both a DNA rearrangement and a local sequence change.

To date, it appears that most of the well-studied microbial strains employ, in parallel, each of the three natural strategies in the generation of genetic variations. In addition, bacteria very often use not only one but also several different specific mechanisms for mutagenesis by each of the strategies. Dissimilar specific mechanisms often function with different efficiencies, as reflected by their contribution to the overall mutation rate. For any given strategy, it might be less relevant as to which specific mechanism is at work than that the particular strategy finds its application with an evolutionarily useful efficiency. In other words, specific mechanisms may substitute for each other within a strategy, at least to some degree. However, this rule does not apply between the strategies because of the difference in the qualities of their evolutionary contributions.

The efficiency displayed by a given specific mechanism of spontaneous mutagenesis may depend on both internal (e.g., the availability of enzymes that mediate mutagenesis) and external factors (e.g., environmental stress). It should also be noted that some mechanisms may act more or less randomly along a DNA molecule, while other mechanisms may demonstrate regional or site preferences for their activities. In view of these considerations, there is a tendency to assume that an evolutionarily fit (or well-prepared) organism should best be able to use a few specific mutagenesis mechanisms for each of the three strategies to generate genetic variations. The meaning of the term ''evolutionary fitness'' will be explained in Section 8.

8
Evolution Genes and Their Own Second-Order Selection

The attentive reader will have seen in the description of some specific mechanisms contributing to the spontaneous mutagenesis that, besides a number of nongenetic factors, specific products of genes are very often at work. These gene products can belong to systems for the repair of DNA damage and will, in this case, modulate the frequency of mutagenesis. Similarly, restriction enzymes seriously reduce both the chance of DNA acquisition and the size of a DNA segment that may eventually be acquired by the recipient cell. Other gene products such as transposases and other mediators of DNA recombination act as generators of genetic variations. Since variation generators and modulators of the frequencies of genetic variation are evolutionary functions, the underlying genetic information is referred to as *evolution genes*. In the

microbial world, these evolution genes generally play no essential role in the physiology of individual lives moving by cell division from one generation to the next. Under laboratory conditions, neither restriction-modification systems nor enzymes for DNA rearrangements are required for the propagation of bacteria. The role of these enzyme systems is primarily evolutionary, and this becomes manifest at the level of populations.

It is assumed that evolution genes are themselves also submitted to selective pressure. However, such selection cannot follow the rules of direct selection for improvements of essential functions, such as those of housekeeping genes. Rather, the selection for the presence and improvement of a variation generator will be exerted at the level of populations. Clearly, it will also be an individual that may one day undergo a mutation improving an evolutionary function. That function will also be exerted in its progeny, in which appropriate genetic variants of genes for directly selected products will be either more or less abundant. Any genetic alteration that affects an evolution gene and proves at longer-term of higher evolutionary value will be maintained, and will provide to the carrier of the involved gene an evolutionary benefit. In the long run this will lead to a fine-tuning of the evolutionary functions of both variation generators and of modulators of the frequency of genetic variation. The underlying indirect selection based on the cells ability to provide genetic variants at a well-balanced level is termed *second-order selection* [26].

It is important to be aware that some gene products may exert their essential functions for the benefit of both the life of the individual and the evolutionary progress of the population. In these cases, it is assumed that evolutionary selection is exerted for both types of function, and will eventually bring the gene to a fine-tuned state to carry out its functions for each of the different purposes on an optional basis. However, as noted above, a number of gene products involved in genetic variation are inessential for the lives of individual bacterial cells. Similarly, the products of many housekeeping genes are inessential for biological evolution.

9 Arguments for a General Relevance of the Theory of Molecular Evolution for All Living Organisms

Largely based on evidence from microbial genetics, it has so far been postulated that the products of a number of evolution genes contribute each in their specific way to the generation of genetic variants at evolutionarily useful frequencies. Thereby, the sources of mutagenesis may relate either to the activity of the evolution gene product itself (e.g., a transposase) or to a nongenetic factor (e.g., a chemical mutagen or an intrinsic structural flexibility of a nucleotide). In many cases, nongenetic factors and the products of evolution genes cooperate in the formation of genetic variants at physiologically tolerable and evolutionarily beneficial levels. This is, for example, the case in spontaneous mutagenesis by an environmental mutagen, when some of the primary damage on the DNA is successfully repaired while other damage may lead to a fixed mutation.

The theory of evolution postulates that life on Earth started almost four billion years ago with primitive, unicellular microorganisms. It was during the first two billion years that microbes must have

developed the basis for the actual set-up of evolutionary strategies and the underlying evolution genes. It can be postulated that the acquired evolutionary capacities could have allowed some microbial populations to undergo a division of labor in increasingly stable associations of cells, and this development might later have led to multicellular organisms. In this type of development, the evolutionary fitness of the involved organisms might have been an important precondition. At still later stages of further evolutionary development, the three natural strategies for generating genetic variations (see Sects 4–7) must have continued exerting their evolutionary influence, together with some other factors such as the formation of symbiotic associations [27]. As a matter of fact, considerable evolutionary relevance is attached to the endosymbiosis of higher organisms with bacteria. Such situations of cohabitation may form an ideal condition for occasional horizontal gene transfer between the close associates.

Clearly, there is still need to conduct research into the spontaneous generation of genetic variation in higher organisms, ideally at the level of the genomes. While this is already quite difficult in microorganisms, it is of increased perplexity with the much larger genomes of higher organisms. However, sequence comparisons today offer fruitful ways of searching for sequence homologies, sequence similarities and single nucleotide polymorphisms, as well as for traces of intragenomic DNA rearrangements and of the horizontal transfer of genetic information. The data available to date are in support of the principles of molecular evolution outlined in Section 2, which are likely to be valid for any type of living organism.

Some of the general evolutionary strategies developed in microorganisms must, ultimately, have also been very useful for the developmental and physiological processes at somatic levels of higher organisms. The generation of antibody diversity in the immune systems of vertebrates by genetic rearrangements, and so-called somatic mutagenesis, is a good example. Another example is the enzymatic repair of DNA damage caused in somatic cells by external mutagens such as ultraviolet (UV) irradiation.

These considerations illustrate that, whichever gene function may prove to be useful for whatever particular purpose, it may be evolutionarily retained and, during the course of time, further fine-tuned. This principle has already been encountered in the microbial world, where the presence of multifunctional enzymes (i.e., those working both for the physiology of the cells and for an evolutionary task) has been postulated to be evolutionarily improved, both by direct and second-order selection for the various functions. Indeed, this may also be the case in higher organisms.

10 Systemic Aspects of Biological and Terrestrial Evolution

As noted above, the updated knowledge on molecular Darwinism clearly indicates that biological evolution is a systemic process of enormous dimensions [25], and it is possible to trace the different types of more or less related organisms back to their common origin. Moreover, in their future evolutionary development, any organism may – at one time or another – also benefit from the acquisition of genetic information from another type of organism by horizontal gene transfer, as has already occurred during the past

evolution. Hence, living organisms not only have a common past – as previously postulated by Charles Darwin – but also a common future, at least to some degree.

Natural selection is a highly complex, permanent process, and substrates for such selection are by far not limited to genetically determined phenotypic traits. Rather, it is becoming apparent that phenotypic traits can become modulated by epigenetic impacts which may, in turn, depend to a large extent on environmental conditions.

In any ecosystem, many quite different organisms cohabit and can mutually influence each other enormously. Most often, such interdependencies take the form of symbiosis, where the different types of organism involved generally help each other, so that all participants can benefit. These aspects play their role in worldwide natural selection. Only rarely does such cohabitation bring benefit to only one of the interacting partners, and in extreme cases the second partner may suffer from a close interaction. This is the case for pathogenesis, which may be regarded as a form of negative selection for the suffering partner.

Today, it is clear that in the long term, the inanimate world is in no way absolutely stable. Both, cosmic and terrestrial evolutions affect the terrestrial living conditions to some degree, and this in turn exerts an influence on habitats and also on the ecological communities in these habitats.

During the course of their cultural evolution, human beings have learned to counteract, to some extent, certain influences of negative natural selection. As is the case for any living organism, humans form an integral part of the worldwide system of evolving communities of forms of life in a developing world. A deeper scientific insight and comprehension of the process of evolution can serve as guidance for future cultural evolution if there is a tendency to avoid causing disturbances to the inherited treasure of the evolutionary pathway. It is important to be aware that the natural processes of evolution in this relatively stable world occur extremely slowly, but steadily. But, this is not a reason to ignore evolutionary constraints; rather, the rate of evolution should serve as guidance for the responsible, sustainable usage of the natural diversities of forms of life and of habitats.

11
Conceptual Aspects of the Theory of Molecular Evolution

With the information provided in Section 2.2 in mind it is fair to re-state that, for the time being, evolution genes and evolution functions represent a concept rather than being a fully proven fact. This concept is based on a particular approach to interpreting the numerous experimental data available. In the following subsections, a brief analysis will first be provided of the difficulties encountered when clarifying the situation in a scientific debate. This will be followed by highlighting the philosophical and more practical values of a deeper understanding of the molecular processes that drive biological evolution.

11.1
Pertinent Scientific Questions

In the history of scientific investigations, biologists have often sought evidence that living organisms could specifically modify, or adapt, their genetic information

in order to better cope with upcoming changes in their living conditions. Most of these attempts have failed to provide the expected response. In some cases – where a certain degree of specific adaptation was observed – specific causal explanations have occasionally been found following more extensive investigations. There is at present, however, no good scientific evidence for a general rule that genetic alterations would always be directed towards a specific goal. This situation favors the view that spontaneous mutations affect DNA more or less randomly, which is in line with the general observation that only a minority of spontaneous mutants will prove to be favorable under the living conditions encountered.

The postulation that evolution genes can act as generators of genetic variations may be surprising in this context, since it is not consistent with the widely held concept of genetic information being a strict program for the fulfillment of a specific task. Whilst this definition does apply to many housekeeping genes (the products of which efficiently catalyze a reaction that reliably yields always the same product), it is not the way that a variation generator functions – inefficiently and yielding a different product from case to case (a good example of this is the transposition of mobile genetic elements). Today, not all scientists see in genetic variation the primary function of a mobile genetic element; rather, some interpret IS translocation – which often goes hand-in-hand with replication of the element – as a selfish activity. In fact, this interpretation considers mobile genetic elements as parasites, with the argument that their activity would often harm the host cell.

This discussion shows that the concept of evolution genes cannot easily be defended by referring to scientific evidence; the concept reflects rather an attitude of the observer of natural events. According to the view defended in this chapter, Nature actively cares for biological evolution. The products of evolution genes are actively involved in generating different types of genetic variants at frequencies which insure not only a genetic stability that is required to maintain the concerned form of life, but also a low frequency of genetic variations as the driving force of evolution. This interpretation recognizes biological evolution as an essential principle of self-organization of life on Earth. In other words, the living world benefits from the natural potency to evolve at any level of life-form.

Another pertinent question with no easy scientific answer relates to the evolutionary function of viruses. Some viruses are clearly identified to serve (on occasion) as gene vectors in horizontal gene transfer, and some also temporarily integrate their genome into the host genome. This situation relates to the lysogenic state of bacteria as well as to endogenous viruses such as retroviruses that reside in many higher organisms. Again, it might be questioned whether primarily, these viruses fulfill evolutionary functions for the evolutionary development of their hosts, or whether they should rather be regarded as parasites, which may carry out some evolutionary function by accident.

While prokaryotic organisms have genomes which are relatively densely packed with functional genes, many higher organisms have extended segments of intergenic DNA sequences, some of which are highly repetitive. Some of these noncoding sequences are highly homogeneous with regard to their nucleotide composition. While the biological roles played by noncoding regions are still

poorly understood, it has been postulated that compositional constraints may influence natural selection [28, 29]. These aspects, which have not been detailed in this chapter, may more specifically relate to the molecular evolution of higher organisms, in addition to the principles outlined here, largely on the basis of evidence from microbial genetics.

11.2 Philosophical Values of the Knowledge on Molecular Evolution

One of the central questions of human curiosity is to know where life – and, more specifically, human life – comes from. The Darwinian Theory opposed the idea of a specific act of creation for each particular form of life, but rather proposed an alternative explanation of a steady evolutionary development which implied the descent of actual species from common ancestors. Hence, the directions of evolution are provided by natural selection acting steadily on all available forms of life, including all present variants. Until recently, there was no specific scientific explanation as to how genetic variants are generated, but with recently developed research strategies it has become possible for molecular genetics to fill this gap. The branch of science termed *molecular evolution* explains that there is not just a single source for genetic variants; rather, many different specific mechanisms contribute to the generation of genetic variants at low frequencies. These mechanisms follow one – and sometimes more than one – basic strategy of evolutionary development. These include local changes in the DNA sequences, rearrangements of the DNA segments within the genome, and the acquisition of segments of foreign DNA by horizontal gene transfer. As a general rule, spontaneous mutagenesis is not specifically directed but is – at least to some extent – a random process, so that only a minor fraction of spontaneous genetic variants will be favorable for the organism concerned and thus provide it with a selective advantage. Nevertheless, new knowledge on the precise molecular mechanisms of the generation of genetic variations provides strong evidence that, in many cases, specific enzymes are involved: these are the products of evolution genes. Such products operate in tight collaboration with nongenetic factors, which may be either intrinsic properties of matter or environmental conditions. This view of the evolutionary process represents the core of a theory of molecular evolution, and it can be seen as an extension of neo-Darwinism to the molecular level (for further detail, see Refs [30–35]).

It should be clearly stated that the theory of molecular evolution does not explain the origin of life. It can, however, explain that biological evolution exerted in all living beings is a steady, dynamic process that is actively promoted not only by intrinsic properties of matter, but also by the intervention of products of evolution genes or more generally of evolutionary functions of many different gene products. A recently published book devoted to these exciting insights is entitled *Darwin in the Genome* [36].

The high philosophical value of this extension of our world view is obvious, and merits extensive discussion and evaluation with regards to its various cultural dimensions. One interesting aspect is the implied duality of the genome. Indeed, evolution genes are located together with all the other genes on the genome, and also on accessory DNA molecules such as plasmids and viral genomes. Whilst it is most likely that a major part of the

gene products carries out functions that will benefit the cell and the individual (often a multicellular organism), it is also probable that a minority of gene products will be involved in the biological evolution of the population concerned. It should be noted, however, that the generation of a novel genetic variant clearly also occurs in an individual cell. Yet, this act of creation has only a small chance of bringing to the cell concerned (and to its descendants) a selective advantage. More often, the mutation will be unfavorable and render the life of the organism more troublesome. However, while the spontaneous rate of mutagenesis remains low (mostly due to an intervention of the evolution genes that are fine-tuned for their activities), unfavorable mutations are tolerable at the level of propagating populations.

Incidentally, the situation described here offers a possible explanation to the rather difficult theodicean question: ''Why does God, despite His love for the human creature, admit that physically evil events such as a mutation causing an inheritable disease can occur to individuals?'' As a matter of fact, in Genesis the Creation is described as a stepwise process, and the descendants of Adam and Eve are not considered to be clones. This implies the permanent evolutionary expansion of the diversity of life-forms [37]. It is also stated in Genesis that God evaluated this system as ''good,'' and hence from a Biblical point of view biological evolution occurs according to God's intention to amplify life diversity on planet Earth. Occasional unfavorable mutations that affect rare individuals among populations represent a sacrifice brought to the creative force residing in the natural potency for molecular evolution. Creation is, therefore, a permanent process.

In brief, the genetic information contained in each genome – of bacteria as well as of all higher organisms – represents an internal duality. It serves individuals for the fulfillment of their individual lives, and it serves populations for a slow but steady expansion of life-forms, and thus of biodiversity.

11.3 Aspects Relating to Practical Applications of Scientific Knowledge on Molecular Evolution

Today, living organisms occupy an amazing variety of ecological niches on planet Earth; these include extreme physico-chemical conditions such as elevated temperatures, high pressure, and also quite unusual compositions of chemical elements. Yet, despite the intrinsic potency of the living world to expand on an evolutionary basis, the carrying capacity of Earth for life can be estimated as in the order of 10^{30} living cells. Whilst this is indeed a very large number, it seriously limits the free expansion of life in its various forms.

The following reflections should help to illustrate this statement. An adult human being carries in the order of 10^{13} cells; consequently, today's human population occupies about 10^{23} cells of the available 10^{30}. In fact, this is close to the average available to each of the estimated 10^{7} different species of organisms on Earth. By extending the reflections made in Section 3, regarding the propagation of bacteria by cell division, it can be concluded that an inoculum of a single bacterial cell would, in theory, lead to the production of 10^{30} cells within only 50 h. In reality, however, growth will be halted at a much earlier stage by a lack of nutrition. Nonetheless, this reflection

illustrates very well the enormous internal forces involved in the expansion of a given form of life.

Similarly, a high potency for evolutionary expansion toward more diversity resides in the mechanisms of molecular evolution that are described in this chapter. These mechanisms can serve us as a basis to better understand both the origin and the steady – though slow – replenishment of biodiversity, as well as the internal limits set to the evolutionary expansion. This knowledge can – and should – be used increasingly as a background for any measures taken in view of the protection of biodiversity and of habitats for diverse forms of life. Not least, a better understanding of the evolutionary process may be helpful to render the development of agricultural and related practices more sustainable.

Genetic engineering offers ample new, sustainable possibilities to produce medicinal drugs, to provide foods of higher quality, and to reduce the nocive impact of the human civilization on the environment [38]. Serious reservations made by large parts of the human population impede many of the proposed biotechnological applications. A part of these concerns refers to the unpredictable long-term effects of genetically modified organisms (GMOs) that are released into the environment, such as in agricultural applications. Hence, scientific assessments of long-term and in particular evolutionary effects of such applications are required, and knowledge of the natural strategies of molecular evolution can provide a good basis for such investigations [39]. In fact, in genetic engineering DNA sequence alterations are brought about within the genome by site-directed mutagenesis in studies of the biological functions of specific genes. In addition, well-defined segments of DNA are introduced into other organisms, with a view either to amplify a particular DNA segment or to harvest a specific gene product. GMOs can also directly serve in applications as genetically modified food, and for bioremediation by microorganisms. Yet, a candid comparison of the practices involving genetic engineering with the natural strategies of generation of genetic variations reveals a high degree of similarity. Typically, the amounts of nucleotides or lengths of DNA sequences involved in these genetic modifications – both in genetic engineering and in natural genetic variation – are of the same order of magnitude. Depending on the strategy involved, they may concern one to a few base pairs, or in other instances a DNA segment containing a sequence domain or one to a few genes, both in intragenomic DNA rearrangements and in the horizontal transfer of DNA between two different organisms. Thus, it can be principally expected that the long-term evolutionary risks of GMOs compare with biohazards intrinsic to the natural process of biological evolution. Similar risks may also be inherent to classical breeding techniques.

These considerations ask for a more integral, holistic, and critical evaluation of the impact of past and present human activities on the natural process of biological evolution. Such assessments should address any human impact on genetic variation, natural selection, and isolation. Based on a long-term historical knowledge, the foundations of life and its evolutionary development on Earth are relatively robust. However, whilst this is good news for humankind, attention should not be deflected from maintaining a responsible and well-intentioned use of scientific knowledge in attempts to create a more easy and comfortable lifestyle.

References

1 Lederberg, J. (1947) Gene recombination and linked segregation in *E. coli*. *Genetics*, **32**, 505–525.
2 Luria, S.E., Delbrück, M. (1943) Mutations of bacteria from virus sensitivity to virus resistance. *Genetics*, **28**, 491–511.
3 Avery, O.T., MacLeod, C.M., McCarty, M. (1944) Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from *Pneumococcus* type III. *J. Exp. Med.*, **79**, 137–158.
4 Kucherlapati, R., Smith, G.R. (Eds) (1988) *Genetic Recombination*, American Society for Microbiology, Washington, DC.
5 Arber, W. (1965) Host-controlled modification of bacteriophage. *Annu. Rev. Microbiol.*, **19**, 365–378.
6 Shapiro, J.A. (1983) *Mobile Genetic Elements*, Academic Press, New York.
7 Watson, J.D., Crick, F.H.C. (1953) Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature*, **171**, 737–738.
8 Drake, J.W. (1991) Spontaneous mutation. *Annu. Rev. Genet.*, **25**, 125–146.
9 Echols, H., Goodman, M.F. (1991) Fidelity mechanisms in DNA replication. *Annu. Rev. Biochem.*, **60**, 477–511.
10 Watson, J.D., Crick, F.H.C. (1953) The structure of DNA. *Cold Spring Harbor Symp. Quant. Biol.*, **18**, 123–131.
11 West, S.C. (1992) Enzymes and molecular mechanisms of genetic recombination. *Annu. Rev. Biochem.*, **61**, 603–640.
12 Glasgow, A.C., Hughes, K.T., Simon, M.I. (1989) Bacterial DNA Inversion Systems, in: Berg, D.E., Howe, M.M. (Eds) *Mobile DNA*, American Society for Microbiology, Washington, DC, pp. 637–659.
13 Sandmeier, H., Iida, S., Arber, W. (1992) DNA inversion region Min of plasmid p15B and Cin of bacteriophage P1: evolution of bacteriophage tail fiber genes. *J. Bacteriol.*, **174**, 3936–3944.
14 Sandmeier, H. (1994) Acquisition and rearrangement of sequence motifs in the evolution of bacteriophage tail fibers. *Mol. Microbiol.*, **12**, 343–350.
15 Iida, S., Hiestand-Nauer, R. (1987) Role of the central dinucleotide at the crossover sites for the selection of quasi sites in DNA inversion mediated by the site-specific Cin recombinase of phage P1. *Mol. Gen. Genet.*, **208**, 464–468.
16 Arber, W. (1991) Elements in microbial evolution. *J. Mol. Evol.*, **33**, 4–12.
17 Rozsa, F.W., Viollier, P., Fussenegger, M., Hiestand-Nauer, R.M., Arber, W. (1995) Cin-mediated recombination at secondary crossover sites on the *Escherichia coli* chromosome. *J. Bacteriol.*, **177**, 1159–1168.
18 Galas, D.J., Chandler, M. (1989) Bacterial Insertion Sequences, in: Berg, D.E., Howe, M.M. (Eds) *Mobile DNA*, American Society for Microbiology, Washington, DC, pp. 109–162.
19 Naas, T., Blot, M., Fitch, W.M., Arber, W. (1994) Insertion sequence-related genetic variation in resting *Escherichia coli* K-12. *Genetics*, **136**, 721–730.
20 Arber, W., Naas, T., Blot, M. (1994) Generation of genetic diversity by DNA re-arrangements in resting bacteria. *FEMS Microbiol. Ecol.*, **15**, 5–14.
21 Naas, T., Blot, M., Fitch, W.M., Arber, W. (1995) Dynamics of IS-related genetic rearrangements in resting *Escherichia coli* K-12. *Mol. Biol. Evol.*, **12**, 198–207.
22 Sengstag, C., Arber, W. (1983) IS2 insertion is a major cause of spontaneous mutagenesis of the bacteriophage P1: non-random distribution of target sites. *EMBO J.*, **2**, 67–71.
23 Lorenz, M.G., Wackernagel, W. (1994) Bacterial gene transfer by natural genetic transformation in the environment. *Microbiol. Rev.*, **58**, 563–602.
24 Arber, W. (2006) The evolutionary strategy of DNA acquisition as a possible reason for a universal genetic code. *Hist. Philos. Life Sci.*, **28**, 525–532.
25 Arber, W. (2009) Systemic aspects of biological evolution. *J. Biotechnol.*, **144**, 242–244.
26 Weber, M. (1996) Evolutionary plasticity in prokaryotes: a panglossian view. *Biol. Philos.*, **11**, 67–88.
27 Magulis, L. (1981) *Symbiosis in Cell Evolution: Life and its Environment on the Early Earth*, Freeman, San Francisco, CA.
28 Bernardi, G. (2000) The compositional evolution of vertebrate genomes. *Gene*, **259**, 31–43.

29 Bernardi, G. (2007) The neo-selectionist theory of genome evolution. *Proc. Natl Acad. Sci. USA*, **104**, 8385–8390.
30 Arber, W. (1993) Evolution of prokaryotic genomes. *Gene*, **135**, 49–56.
31 Arber, W. (1995) The generation of variation in bacterial genomes. *J. Mol. Evol.*, **40**, 7–12.
32 Arber, W. (2000) Genetic variation: molecular mechanisms and impact on microbial evolution. *FEMS Microbiol. Rev.*, **24**, 1–7.
33 Arber, W. (2002) Evolution of prokaryotic genomes. *Curr. Top. Microbiol. Immunol.*, **264/I**, 1–14.
34 Arber, W. (2003) Elements for a theory of molecular evolution. *Gene*, **317**, 3–11.
35 Caporale, L.H. (Ed.) (1999) Molecular Strategies in Biological Evolution. *Ann. N. Y. Acad. Sci.*, **870** (Special issue).
36 Caporale, L.H. (2003) *Darwin in the Genome: Molecular Strategies in Biological Evolution*, McGraw-Hill, New York.
37 Arber, W. (2007) Sagesse Traditionelle et Savoir Récemment Acquis sur L'évolution Biologique, in: Staune, J. (Ed.) *La Science, L'homme et le Monde – les Nouveaux Enjeux*, Presses de la Renaissance, Paris, pp. 127–149.
38 Arber, W. (2009) The impact of science and technology on the civilization. *Biotechnol. Adv.*, **27**, 940–944.
39 Arber, W. (2002) Molecular evolution: comparison of natural and engineered variations. *Pontif. Acad. Sci. Scr. Varia*, **103**, 90–101.